

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 07-210330

(43)Date of publication of application : 11.08.1995

(51)Int.Cl. G06F 3/06
G06F 3/06

(21)Application number : 06-001980

(71)Applicant : FUJI XEROX CO LTD

(22)Date of filing : 13.01.1994

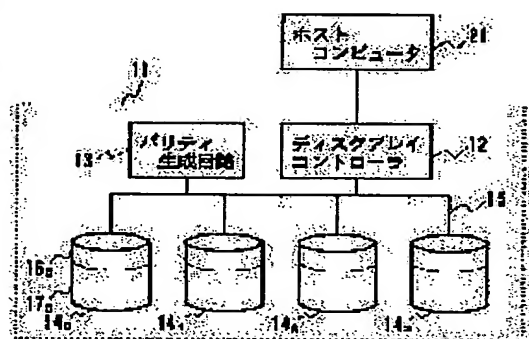
(72)Inventor : OIZUMI SHIGERU

(54) DISK ARRAY DEVICE

(57)Abstract:

PURPOSE: To respond to a write request at a high speed and to complete the processing at a high speed by distributing and storing data for which write is requested in a buffer area with a storage area capable of transferring the data at a high speed as a buffer and reconstituting the storage form of the data to redundant constitution when no access request is present.

CONSTITUTION: The data for which the write is requested are tentatively written in the free area of a buffer area 160 which is the storage area whose transfer rate is high inside respective disk devices 140-143 and the data for which the write is requested are divided into the blocks of a prescribed size and distributed and stored in the four disk devices 140-143. Then, when no access request from an outside is present, a disk array controller 12 transfers the data stored in the buffer area 160 to the storage area 170 under conditions that the disk devices 140-143 of a transfer origin and a transfer destination are different, a parity is generated by using a parity generation circuit 13 and the storage form of the data is reformed to the redundant constitution.



LEGAL STATUS

[Date of request for examination] 18.09.2000

[Date of sending the examiner's decision of rejection] 27.04.2004

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

Copyright (C); 1998,2003 Japan Patent Office

* NOTICES *

JPO and NCIPi are not responsible for any damages caused by the use of this translation.

- 1.This document has been translated by computer. So the translation may not reflect the original precisely.
- 2.**** shows the word which can not be translated.
- 3.In the drawings, any words are not translated.

CLAIMS

[Claim(s)]

[Claim 1] Two or more sets of disk units and the buffer area used for the temporary storage of data to said field which can perform two or more data transfer at the high speed in each storage region of the disk unit of a base are created. A creation means to create the data area used for storing of data made into the redundant configuration in the remaining storage regions, A division means to divide into the data block of predetermined size the data with which the write request was made, The data block divided by this division means Said write-in means written in the buffer area of the disk unit of a base, [two or more] A judgment means said to judge whether two or more data blocks of sufficient amount for calculation of parity are stored in the buffer area of the disk unit of a base, When said judgment means judges with the data block of sufficient amount for calculation of parity being stored A data transfer means to transmit the data block stored in the buffer area to the data area of the same address in a different disk unit from the disk unit stored in the data block, Said data transfer means creates parity based on the data block transmitted to the data area. Disk array equipment characterized by providing the parity write-in means which writes the created parity in the data area specified in said address of the disk unit which is not used for storing of those data blocks.

[Claim 2] Two or more sets of disk units, and a creation means to create said data area of the predetermined number in each storage region of the disk unit of a base corresponding to the data transfer rate two or more, A detection means to detect the compressibility of the data with which the write request was made, and a selection means to choose from the data area of said predetermined number the data area used for storing of that data according to the compressibility which this detection means detected, Disk array equipment characterized by providing the write-in means which divides and writes said data in the block of predetermined size in the data area chosen by this selection means.

[Translation done.]

* NOTICES *

JPO and NCIPi are not responsible for any damages caused by the use of this translation.

- 1.This document has been translated by computer. So the translation may not reflect the original precisely.
- 2.**** shows the word which can not be translated.
- 3.In the drawings, any words are not translated.

DETAILED DESCRIPTION

[Detailed Description of the Invention]

[0001]

[Industrial Application] This invention relates to the disk array equipment using the transfer rates of a disk unit differing with respect to disk array equipment according to the storage region where data are memorized.

[0002]

[Description of the Prior Art] Compared with storage elements, such as semiconductor memory, the disk unit used as external storage, such as a computer, can memorize mass data, and has flume *****. However, since mechanical actuation is needed for read-out of data, it also has the fault that a speed of response is slow or the information memorized on disk media by the collision of a head and disk media etc. may be lost. For this reason, improvement in a speed of response or dependability is aimed at, data are divided into the block of predetermined size, and storing in two or more sets of disk units dispersedly is performed.

[0003] Although the data storage of such a form is realizable also by publishing a command from a computer side to each of two or more sets of disk units, the burden by the side of a computer will become large. In order to solve this problem, the equipment which opted for the storing procedure by the store side is disk array equipment.

[0004] To the storing gestalt well used with disk array equipment, it is RAID (Redundant Arrays of Inexpensive Disks). There is a redundant configuration called. the writing of the parity created based on the writing of data, simultaneously its data with the gestalt classified into level 3 thru/or 5 among these storing gestalten -- a line -- **. With conventional disk array equipment, it has the semiconductor memory for memorizing temporarily the data with which the write request was made, and parity was computed based on the data memorized on the memory.

[0005]

[Problem(s) to be Solved by the Invention] As explained above, with conventional disk array equipment, data were memorized on semiconductor memory temporarily and parity was computed using the data. For this reason, when the write request of a lot of data exceeding the problem that it is required to have the semiconductor memory which has a certain amount of storage capacity, and the storage capacity of semiconductor memory was received, there was a problem that that writing speed will be restricted by the calculation process of parity.

[0006] Then, the purpose of this invention is to offer the disk array equipment which solves the above problems using the difference in the transfer rate of the storage region in each disk unit, for example.

[0007]

[Means for Solving the Problem] Invention according to claim 1 creates the buffer area used for the temporary storage of data in the field which can perform data transfer at the high speed in two or more sets of each storage regions of two or more sets of a disk unit and disk units. A creation means to create the data area used for storing of data made into the redundant configuration in the remaining storage regions, A division means to divide into the data block of predetermined size the data with which the

write request was made, The write-in means which writes the data block divided by this division means in the buffer area of two or more sets of disk units, A judgment means to judge whether the data block of sufficient amount for calculation of parity is stored in the buffer area of two or more sets of disk units, When a judgment means judges with the data block of sufficient amount for calculation of parity being stored A data transfer means to transmit the data block stored in the buffer area to the data area of the same address in a different disk unit from the disk unit stored in the data block, This data transfer means creates parity based on the data block transmitted to the data area, and possesses the parity write-in means which writes the created parity in the data area specified in the address of a disk unit which is not used for storing of those data blocks.

[0008] That is, in invention according to claim 1, the data with which the write request was made are once distributed and stored in these buffer areas, using as a buffer the storage region which can perform data transfer at the high speed in each disk unit. And when the access request of data is not made, the data stored in the buffer area are transmitted under conditions from which the disk unit of the source and the destination differs, and the storing gestalt of data is reconstituted in a redundant configuration. Thus, since parity is not created and a data transfer rate writes the data in an early storage region when the write request of data is made, a high speed can be answered to a write request. Moreover, since a redundant configuration is formed, a high speed can be made to also complete the processing only by the data transfer between different disk units.

[0009] A creation means by which invention according to claim 2 creates the data area of the predetermined number according to a data transfer rate within two or more sets of each storage regions of two or more sets of a disk unit and disk units, A detection means to detect the compressibility of the data with which the write request was made, and a selection means to choose from the data area of a predetermined number the data area used for storing of that data according to the compressibility which this detection means detected, The write-in means which divides and writes data in the block of predetermined size in the data area chosen by this selection means is provided.

[0010] That is, in invention according to claim 2, using the storage region where data transfer rates differ existing in the storage region of a disk unit, the data with bad compressibility are stored in the storage region where a transfer rate is high, and the data with sufficient compressibility are conversely stored in the storage region where a transfer rate is low. For this reason, what data with bad compressibility are stored in the storage region where a transfer rate is low is lost, and the synthetic speed of response as a disk unit can be improved.

[0011]

[Example] Hereafter, this invention is explained to a detail per example.

[0012] The 1st example [0013] The configuration of the disk array equipment by the 1st example of this invention is shown in drawing 1. The disk array equipment 11 of the 1st example consists of four sets of the disk array controller 12, the parity generation circuit 13, and disk units 14, and they are mutually connected by bus 15. A disk controller 12 is the control circuit which consisted of a processor, memory, and its circumference circuit, and the parity generation circuit 13 computes the exclusive OR of the data of a predetermined number with directions of the disk array controller 12.

[0014] Actuation of this disk array equipment is explained to below. Although the disk array equipment of the 1st example is equipment which stores by making data into a redundant configuration, the data with which the write request was made are once written in the free area of the buffer area 16 which is the storage region where the transfer rate in each disk unit 14 is high.

[0015] This writing is performed under control of the disk array controller 12, and the data with which the write request was made are divided into the block of predetermined size, and are distributed and stored in four sets of disk units 14. in addition, on the memory in the disk array controller 12 The address which an external device uses for assignment of the data block in disk array equipment, The address translation table for matching the address within the disk unit the data block is actually remembered to be, and a disk unit is prepared. The disk array controller 12 memorizes having stored the data block in the buffer area 16 by finding out the free area in a buffer area and updating those contents by verifying the contents of this address translation table.

[0016] And when the access request from an external device is not made, the disk array controller 12 transmits the data stored in the buffer area 16 to a storage area 17 in the following procedures.

[0017] The flow of actuation of the disk array controller at the time of the data transfer to a storage area is shown in drawing 2. The disk array controller 12 is the buffer area 160 of four sets of disk units first. Or 163 The number NB of data blocks stored It investigates (step S101) and the data block stored In being under the constant NCNST (step S102; N), it waits to once end processing and to write in new data. In addition, NCNST is a constant depending on the number of the disk unit which constitutes disk array equipment, and the relation of $NCNST = M \times (M - 1)$ is materialized between NCNST and the number M of a disk unit. With the disk array equipment of the 1st example, since four sets of disk units are used, a constant NCNST is set to "12."

[0018] When the data block is stored in the buffer area 16 more than the NCNST individual (step S102; Y), the disk array controller 12 chooses from each buffer area 16 first the data block of the NCNST individual made into a processing object (step S103). Subsequently, the storage region within the disk unit used as the destination of each data block and its disk unit is determined and combined, and the order of a transfer is determined (step S104). Disk unit 143 used at this step in order to memorize parity Out of the storage area of three sets of the disk units to remove, the storage region used as the data transfer between the same disk units is chosen, and it is set as the destination. The detail of this configuration procedure is mentioned later.

[0019] The disk array controller 12 which determined the destination and the order of a transfer of a data block performs a transfer of the data block of a NCNST individual according to it (step S105). Next, the parity generation circuit 13 is controlled and it is a disk unit 140. Or 142 The parity which is the exclusive OR of the data block memorized in the storage area of the corresponding address is made to create, and it is a disk unit 143 about the parity. It is made to write in (step S106). And the contents of the address translation table are rewritten (step S107), and actuation is ended.

[0020] The outline of the processing performed in step S104 is explained using drawing 3 thru/or drawing 5. In addition, drawing 3 R> 3 is drawing for explaining the algorithm used by the flow of drawing 4, drawing 4 is the flow chart having shown the flow of the actuation performed in order that a disk array controller may determine the destination, and drawing 5 is the mimetic diagram showing the relation of the destination and the source which are determined as a result of the actuation.

[0021] First, the outline of the decision procedure of the destination is explained using drawing 3. As mentioned above, the destination is chosen for the disk unit excluding [the data block stored in the buffer area of a certain disk unit] the disk. Moreover, since one data block in the data block specified in the same address in a storage area is used for storage of parity, it should just choose the destination from the data block specified in the same address in a storage area for every data of the "number -1 of disk unit" individual.

[0022] For example, like drawing 3 (a), when the data of "0" to "11" and 12 ***** with a number are stored in the buffer area of a disk unit, it extracts a data block at a time to three numerical orders, and if each three data block is arranged in the same line, it will become like drawing 3 (b), without changing the train (namely, disk unit) in which each data block exists. For example, if this line is rotated so that the part on which the data block of the 2nd line of (b) is not drawn may come to most a right end, as shown in (c), each data is movable to the train different from the original train. By the 3rd line and processing that the 4th line is the same, each data which exists in the line can be arranged in the original train and a different train.

[0023] The arrangement obtained by such conversion fulfills above-mentioned conditions, and, fundamentally, the decision of the destination by the disk array controller explaining the back is made with such an algorithm. However, the destination is determined by the algorithm with which the 1st line of drawing 3 (a) differs from an above-mentioned thing since the part on which the data block is not drawn has already existed in the right end train.

[0024] The processing which a disk array controller actually performs is explained using drawing 4. The block number which this processing asks for the number of the data block in the storage area used as the destination of the data block of No. j in a buffer area, and was obtained is Variable Aj. It

memorizes. In addition, block number j is prescribed by $U_x (L-1)$, U is the number (0-3) of the disk unit in which that data block was stored, that data block $[L]$ in each area is the order of storing, and this block number is independently prescribed by the buffer area and the storage area with the disk array equipment of the 1st example.

[0025] First, a disk array controller sets "0" to Counters NBUFF and NDATA (step S201), and, subsequently to i and j , sets "0" and "NBUFF+UMAX" to them, respectively (step S202). Here, it is UMAX. It is the maximum of a disk unit number and is the value which subtracted "1" from the number of a disk unit which disk array equipment contains. Moreover, Counter NBUFF is a counter for memorizing the block number of the beginning of a series of blocks made into the processing object in a buffer area, and NDATA is a counter for the block number of the beginning of a series of data blocks used as the destination in a storage area to memorize.

[0026] And U_j which is the disk unit number in which the data block of No. j is stored and UMAX It compares and the data block of No. j is UMAX. In not being what was stored in the disk unit of watch (step S203; N), it determines the destination of the data block of "NBUFF+ i " watch as the data block of the "NDATA+ i " watch in a storage area (step S204). And "1" is added to i (step S205), and i is UMAX. If not in agreement (step S206; N), return and i are UMAX to step S204. When in agreement, it progresses to (step S206; Y) and step S211.

[0027] Moreover, UMAX When stored in the disk unit of watch (step S203; Y), the destination of the data block of "NBUFF+ i " watch is determined as the data block of the "NDATA+ $i+1$ " watch in a storage area (step S207). And if "1" is added to i (step S208) and i is not in agreement with "UMAX-1" (step S209; N), it returns to step S207. And when i is in agreement with "UMAX-1", the destination of the data block of (step S209; Y) and " $j+1$ " watch is determined as the data block of the "NDATA" watch in a storage area (step S210), and it progresses to step S211. In addition, the processing currently performed at this step S207 thru/or step S210 is processing to the party eye of drawing 3.

[0028] At step S211, if "UMAX+1" and UMAX are added, respectively and NBUFF is not in agreement with NDATA and NBUFF with NCNST (step S212; N), it returns to step S202. And when NBUFF and NCNST are in agreement, setting processing of (step S212; Y) and the destination is ended.

[0029] By this processing of a series of, as shown in drawing 5, the data block in three sets of the storage areas of a disk unit is specified as each of the data block memorized by the buffer area of four sets of disk units as a storing place. The figure written in each data block in drawing 5 is the data number attached in order to identify data, and is a disk unit 143. PL-M written to the storage area The parity computed based on the data to that of No. L to No. M is shown. In addition, the data number is in agreement with a block number within a buffer area.

[0030] With the disk array equipment of the 1st example, in this way, since the data transfer point is specified as other disk units, actuation of writing in while reading data is attained and data transfer can be completed by short time amount. If the sequence of performing a data transfer is not suitable, a head must stop however, having to seek a long distance after completion of the data transfer of each block. For example, in order of a data number, if data transfer is made to perform, many cases where the head used for writing is used for read-out by the next processing will appear. For this reason, with the disk array equipment of the 1st example, after making the writing of data to each disk unit complete, the procedure of performing the writing to the following disk unit is performing data transfer.

[0031] That is, when performing data transfer to the condition of drawing 5, it is a disk unit 140 first. All the data transfer to an inner ($U=0$) storage area will be made to complete. Before starting a transfer of the last "data 9" at this time, it is a disk unit 142. A head is moved to a storage area side. And disk unit 140 After completion of data transfer, and disk unit 142 Data transfer is made to start.

[0032] In such order, the time amount needed for data transfer is further shortened with the disk array equipment of the 1st example by performing data transfer.

[0033] In addition, although the case where the redundant configuration of the RAID level 4 was made to form was made into the example, data can also be made to memorize by the redundant configuration of the RAID level 5 with the disk array equipment of the 1st example in the above explanation. In this case, if it shifts, every $["1"]$ and the destination which fulfills required conditions can be set to the right

or the left for the condition which showed in drawing 3 (b) to each data. As reference, the flow of the actuation of a disk array controller in the case of forming the redundant configuration of the RAID level 5 in drawing 6 is shown, and arrangement of the data block obtained by drawing 7 as the result is shown.

[0034] Since the flow shown in drawing 6 is similar with drawing 4, it is made to explain only the outline. The block number of the beginning of the buffer area and storage area used as a processing object goes into Counters NBUFF and NDATA, and the destination of the data block from the NBUFF watch stored in the buffer area to "NBUFF+UMAX -1" watch is determined as them in each loop formation (step S302 thru/or S309). The destination of each data block is set as the disk unit on the right of [where each data block was stored] a disk unit (step S303 thru/or step S308). Naturally, the destination of the data block in a right end disk unit is set up in a left end disk unit (step S306).

[0035] Modification [0036] Although actuation of reading the data for 3 blocks, creating parity, and writing in the created parity is required of the disk array equipment of the 1st example, since a parity generation circuit and each disk unit were connected by one bus, a certain amount of time amount is needed at the time of parity creation processing. In order to shorten this parity creation processing time, equipment consisted of disk array equipment of a modification using the parity generation circuit which can process three data to coincidence, and the bus electronic switch.

[0037] The configuration of the disk array equipment of a modification is shown in drawing 8. The disk array equipment 11 of a modification consists of four sets of the disk array controller 12, the parity generation circuit 13, and disk units 14, and a bus electronic switch 18.

[0038] Actuation of the bus electronic switch 18 and the parity generation circuit 13 is explained using drawing 9. Sometimes, the bus electronic switch 18 usually connects each disk unit 14 with the disk array controller 12. And when creation of parity is directed, as shown in drawing 9, internal connection is changed and the path which can perform an exchange of data each disk unit and directly consists of disk array controllers. three sets of and the disk units 140 Or 142 from -- the parity which the data which compute parity are read to coincidence, and those EXCLUSIVE OR operation is performed, and is the result of an operation -- disk unit 143 It writes in.

[0039] Thus, with the disk array equipment of a modification, since the time amount needed at the time of parity calculation decreases compared with the disk array equipment of the 1st example, compaction of the synthetic operating time can be aimed at.

[0040] The 2nd example [0041] The configuration of the disk array equipment of the 2nd example is shown in drawing 10. The disk array equipment of the 2nd example takes in circuit the configuration which added the compression expansion circuit 19 to the disk array equipment of the 1st example. A compression expansion circuit 19 is a circuit which elongates compression of the data sent from the external device of host computer 21 grade, and compressed data stored in the disk unit, and is a circuit which also performs calculation of compressibility to coincidence at the time of compression. With this disk array equipment, after the data with which the high transfer rate field 22 which is a storage region classified in each disk unit according to the transfer rate, the inside transfer rate field 23, and the low transfer rate field 24 were set up, and the write request was made are compressed, they are stored in the transfer rate field according to that compressibility.

[0042] The flow of the actuation at the time of the data storage of the disk array controller in the disk array equipment of the 2nd example is shown in drawing 11. The disk array controller which received the write request of data chooses the storage region which stores data according to the compressibility of the data (step S401 thru/or step S405). By these steps, compressibility (data size before the data size / compression after compression) is a constant C1. To small data, storing in a low transfer rate field is specified (step S402), and compressibility is a constant C1. It is C2 above. An inside transfer rate field is specified as the data of the following (step S404). Moreover, compressibility is a constant C2. A high transfer rate field is specified as a storing field of the above data (step S405).

[0043] Then, it is judged whether sufficient free area for storing of the data is in the specified storage region (step S406), and when there is a free area, data are stored there at (Y) (step S408). When there is no free area in the storage region chosen based on compressibility (step S406; N), it writes in there in

search of the free area of other storage regions (step S407). In addition, an error message is outputted before starting the processing shown here, when there is no availability which memorizes the data in disk array equipment.

[0044] Thus, with this disk array equipment, the data with bad compressibility are stored in the storage region where a transfer rate is high, and the data with sufficient compressibility are conversely stored in the storage region where a transfer rate is low. For this reason, what data with bad compressibility are stored in the storage region where a transfer rate is low can be lost, and the synthetic speed of response as a disk unit can be raised.

[0045] It constitutes from disk array equipment of the 2nd example so that it may compress within equipment in response to the data which are not compressed from an external device, but you may constitute so that the data compressed from the external device and its compressibility may be received. Moreover, you may make it take the configuration using parity like the RAID level 4, and may make it take the configuration which does not use parity like the RAID level 0 in the writing of steps S406 and S407. Moreover, like the disk array equipment of a modification, in case parity is used, you may constitute combining a bus electronic switch so that a parity generation circuit can access the data in each disk unit at coincidence.

[0046]

[Effect of the Invention] As explained above, in order to use for the high speed in two or more disk units as a buffer the storage region which can perform data transfer, by invention according to claim 1, the working speed at the time of receiving a lot of data at a stretch becomes earlier than conventional disk array equipment. Moreover, since it is constituted so that a redundant configuration may be formed only by the data transfer between different disk units, a redundant configuration can be formed in a high speed.

[0047] In invention according to claim 2, the data with bad compressibility are stored in the storage region where a transfer rate is high, and the data with sufficient compressibility are conversely stored in the storage region where a transfer rate is low. For this reason, what data with bad compressibility are stored in the storage region where a transfer rate is low can be lost, and the synthetic speed of response as a disk unit can be raised.

[Translation done.]

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平7-210330

(43) 公開日 平成7年(1995)8月11日

(51) Int.Cl.⁸

G 0 6 F 3/06

識別記号

5 4 0

3 0 5 C

庁内整理番号

F I

技術表示箇所

審査請求 未請求 請求項の数2 OL (全9頁)

(21) 出願番号

特願平6-1980

(22) 出願日

平成6年(1994)1月13日

(71) 出願人 000005496

富士ゼロックス株式会社

東京都港区赤坂三丁目3番5号

(72) 発明者 大泉 繁

埼玉県岩槻市府内3丁目7番1号 富士ゼロックス株式会社岩槻事業所内

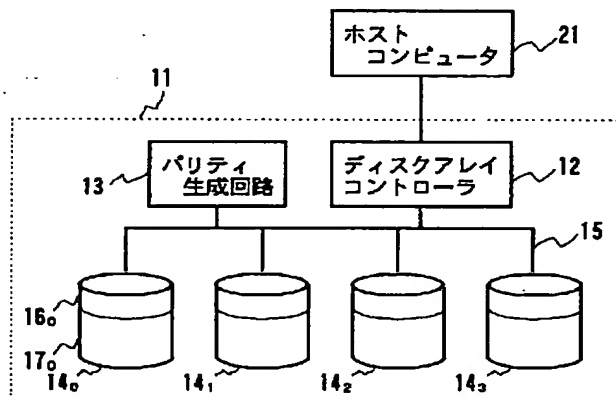
(74) 代理人 弁理士 山内 梅雄

(54) 【発明の名称】 ディスクアレイ装置

(57) 【要約】

【目的】 個々のディスク装置内の転送レートの高い記憶領域をバッファとして利用して、データを冗長構成にして格納するディスクアレイ装置を提供する。

【構成】 それぞれのディスク装置14内の高速にデータ転送が行なえる記憶領域16をバッファとして用い、書き込み要求のなされたデータを、一旦、これらのバッファ領域に分散して格納する。そして、データのアクセス要求がなされていないときに、バッファ領域16に格納したデータを、転送元と転送先のディスク装置が異なるような条件下で、格納領域17に転送するとともに、パリティ生成回路13を用いてパリティを作成し、データの格納形態を冗長構成に形成し直す。



【特許請求の範囲】

【請求項 1】 複数台のディスク装置と、

前記複数台のディスク装置のそれぞれの記憶領域内の高速にデータ転送が行なえる領域にデータの一時記憶に用いるバッファ領域を作成し、残りの記憶領域に冗長構成にしたデータの格納に用いるデータ領域を作成する作成手段と、

書き込み要求がなされたデータを所定サイズのデータブロックに分割する分割手段と、

この分割手段により分割されたデータブロックを前記複数台のディスク装置のバッファ領域に書き込む書込手段と、

前記複数台のディスク装置のバッファ領域にパリティの算出に十分な量のデータブロックが格納されているか否かを判定する判定手段と、

前記判定手段がパリティの算出に十分な量のデータブロックが格納されていると判定したときに、バッファ領域に格納されているデータブロックをそのデータブロックが格納されたディスク装置とは異なるディスク装置内の同一のアドレスのデータ領域に転送するデータ転送手段と、

前記データ転送手段がデータ領域に転送したデータブロックを基にパリティを作成し、作成したパリティをそれらのデータブロックの格納に用いられていないディスク装置の前記アドレスで指定されるデータ領域に書き込むパリティ書込手段とを具備することを特徴とするディスクアレイ装置。

【請求項 2】 複数台のディスク装置と、

前記複数台のディスク装置のそれぞれの記憶領域内に、データの転送速度に応じた所定数のデータ領域を作成する作成手段と、

書き込み要求がなされたデータの圧縮率を検出する検出手段と、

この検出手段が検出した圧縮率に応じて、前記所定数のデータ領域からそのデータの格納に用いるデータ領域を選択する選択手段と、

この選択手段により選択されたデータ領域に、前記データを所定サイズのブロックに分割して書き込む書込手段とを具備することを特徴とするディスクアレイ装置。

【発明の詳細な説明】

【0001】

【産業上の利用分野】本発明は、ディスクアレイ装置に係わり、データが記憶される記憶領域に応じてディスク装置の転送レートが異なることを利用するディスクアレイ装置に関する。

【0002】

【従来の技術】コンピュータなどの外部記憶装置として用いられるディスク装置は、半導体メモリなどの記憶素子に比べて、大容量のデータを記憶することができという利点を有する。しかし、データの読み出しに機械的な

動作を必要とするため、応答速度が遅い、または、ヘッドとディスク媒体の衝突などにより、ディスク媒体上に記憶しておいた情報が失われてしまうことがあるといった欠点も有している。このため、応答速度や信頼性の向上を図って、データを所定のサイズのブロックに分割し、複数台のディスク装置に分散して格納することが行なわれている。

【0003】このような形のデータ格納は、コンピュータ側より、複数台のディスク装置のそれぞれに対してコマンドを発行することによっても実現することはできるが、コンピュータ側の負担が大きくなってしまう。この問題を解消するために、記憶装置側で、格納手順を決定するようにした装置がディスクアレイ装置である。

【0004】ディスクアレイ装置で良く用いられている格納形態に、RAID (Redundant Arrays of Inexpensive Disks) と呼ばれる冗長構成がある。これらの格納形態のうち、たとえば、レベル 3 ないし 5 に分類される形態では、データの書き込みと同時に、そのデータを基に作成されたパリティの書き込みが行なれる。従来のディスクアレイ装置では、書き込み要求がなされたデータを一時的に記憶するための半導体メモリが備えられており、そのメモリ上に記憶されたデータを基にパリティが算出されていた。

【0005】

【発明が解決しようとする課題】以上説明したように、従来のディスクアレイ装置では、データを一時、半導体メモリ上に記憶して、そのデータを用いてパリティの算出を行っていた。このため、ある程度の記憶容量を有する半導体メモリを備えることが必要であるという問題と、半導体メモリの記憶容量を越える大量のデータの書き込み要求を受信したときには、パリティの算出過程によりその書込速度が制限されてしまうという問題があった。

【0006】そこで本発明の目的は、たとえば、上述のような問題を、個々のディスク装置内の記憶領域の転送レートの違いを利用して解消するディスクアレイ装置を提供することにある。

【0007】

【課題を解決するための手段】請求項 1 記載の発明は、複数台のディスク装置と、複数台のディスク装置のそれぞれの記憶領域内の高速にデータ転送が行なえる領域にデータの一時記憶に用いるバッファ領域を作成し、残りの記憶領域に冗長構成にしたデータの格納に用いるデータ領域を作成する作成手段と、書き込み要求がなされたデータを所定サイズのデータブロックに分割する分割手段と、この分割手段により分割されたデータブロックを複数台のディスク装置のバッファ領域に書き込む書込手段と、複数台のディスク装置のバッファ領域にパリティの算出に十分な量のデータブロックが格納されているか否かを判定する判定手段と、判定手段がパリティの算出

に十分な量のデータブロックが格納されていると判定したときに、バッファ領域に格納されているデータブロックをそのデータブロックが格納されたディスク装置とは異なるディスク装置内の同一のアドレスのデータ領域に転送するデータ転送手段と、このデータ転送手段がデータ領域に転送したデータブロックを基にパリティを作成し、作成したパリティをそれらのデータブロックの格納に用いられていないディスク装置のアドレスで指定されるデータ領域に書き込むパリティ書込手段とを具備する。

【0008】すなわち、請求項1記載の発明では、それぞれのディスク装置内の高速にデータ転送が行なえる記憶領域をバッファとして用い、書き込み要求のなされたデータを、一旦、これらのバッファ領域に分散して格納する。そして、データのアクセス要求がなされていないときに、バッファ領域に格納したデータを、転送元と転送先のディスク装置が異なるような条件下で転送していき、冗長構成にデータの格納形態を構成し直す。このように、データの書き込み要求がなされた際には、パリティの作成を行わず、また、そのデータをデータ転送速度が早い記憶領域に書き込むため、書き込み要求に対して高速に応答することができる。また、異なるディスク装置間のデータ転送だけで、冗長構成を形成するので、その処理も高速に完了させることができる。

【0009】請求項2記載の発明は、複数台のディスク装置と、複数台のディスク装置のそれぞれの記憶領域内に、データの転送速度に応じた所定数のデータ領域を作成する作成手段と、書き込み要求がなされたデータの圧縮率を検出する検出手段と、この検出手段が検出した圧縮率に応じて、所定数のデータ領域からそのデータの格納に用いるデータ領域を選択する選択手段と、この選択手段により選択されたデータ領域に、データを所定サイズのブロックに分割して書き込む書込手段とを具備する。

【0010】すなわち、請求項2記載の発明では、ディスク装置の記憶領域にデータ転送速度が異なる記憶領域が存在することを利用して、圧縮率が悪いデータは、転送レートが高い記憶領域に格納し、逆に、圧縮率が良いデータは、転送レートが低い記憶領域に格納する。このため、圧縮率が悪いデータが転送レートの低い記憶領域に格納されるようなことがなくなり、ディスク装置としての総合的な応答速度を向上することができる。

【0011】

【実施例】以下、実施例につき本発明を詳細に説明する。

【0012】第1の実施例

【0013】図1に、本発明の第1の実施例によるディスクアレイ装置の構成を示す。第1の実施例のディスクアレイ装置11は、ディスクアレイコントローラ12とパリティ生成回路13と4台のディスク装置14とで構

成され、それらは、バス15で相互に接続されている。ディスクコントローラ12は、プロセッサとメモリとその周辺回路で構成された制御回路であり、パリティ生成回路13は、ディスクアレイコントローラ12の指示により、所定数のデータの排他的論理和を算出する。

【0014】以下に、このディスクアレイ装置の動作の説明を行なう。第1の実施例のディスクアレイ装置は、データを冗長構成にして格納する装置ではあるが、書き込み要求がなされたデータは、一旦、各ディスク装置14内の転送レートが高い記憶領域であるバッファエリア16の空き領域に書き込まれる。

【0015】この書き込みはディスクアレイコントローラ12の制御下で行われ、書き込み要求がなされたデータは、所定のサイズのブロックに分割されて、4台のディスク装置14に分散されて格納される。なお、ディスクアレイコントローラ12内のメモリ上には、外部装置がディスクアレイ装置内のデータブロックの指定に用いるアドレスと、そのデータブロックが実際に記憶されているディスク装置およびディスク装置内でのアドレスとを対応付けるためのアドレス変換テーブルが設けられており、ディスクアレイコントローラ12は、このアドレス変換テーブルの内容を検証することにより、バッファエリア内の空き領域を見だし、その内容を更新することにより、バッファエリア16に、データブロックを格納したことを記憶する。

【0016】そして、外部装置からのアクセス要求がなされていないときに、ディスクアレイコントローラ12は、バッファエリア16に格納したデータを、以下の手順で格納エリア17に転送する。

【0017】図2に、格納エリアへのデータ転送時における、ディスクアレイコントローラの動作の流れを示す。ディスクアレイコントローラ12は、まず、4台のディスク装置のバッファエリア16₀ないし16₃に、格納されているデータブロック数N_Bを調べ(ステップS101)、格納されているデータブロックが、定数N_{CNST}未満である場合(ステップS102; N)には、処理を一旦終了して、新たなデータが書き込まれるのを待つ。なお、N_{CNST}は、ディスクアレイ装置を構成するディスク装置の台数に依存する定数であり、N_{CNST}とディスク装置台数Mとの間には、N_{CNST} = M × (M - 1)の関係が成立する。第1の実施例のディスクアレイ装置では、4台のディスク装置を用いているので、定数N_{CNST}は、“12”となる。

【0018】バッファエリア16にデータブロックが、N_{CNST}個以上格納されていた場合(ステップS102; Y)、ディスクアレイコントローラ12は、まず、それぞれのバッファエリア16から、処理対象とするN_{CNST}個のデータブロックを選択(ステップS103)する。次いで、それぞれのデータブロックの転送先となるディスク装置とそのディスク装置内での記憶領域を決定し、

併せて、その転送順を決定(ステップS104)する。このステップでは、パリティを記憶するために使用するディスク装置14₃を除く3台のディスク装置の格納エリア内から、同一のディスク装置間のデータ転送とならない記憶領域が選択され、転送先に設定される。この設定手順の詳細は後述する。

【0019】データブロックの転送先および転送順を決定したディスクアレイコントローラ12は、それに従い、N_{CST}個のデータブロックの転送を実行(ステップS105)する。次に、パリティ生成回路13を制御して、ディスク装置14₀ないし14₂の対応するアドレスの格納エリアに記憶されたデータブロックの排他的論理和であるパリティを作成させ、そのパリティを、ディスク装置14₃に書き込ませる(ステップS106)。そして、アドレス変換テーブルの内容を書き換えて(ステップS107)、動作を終了する。

【0020】図3ないし図5を用いて、ステップS104において行なわれる処理の概要を説明する。なお、図3は、図4の流れで用いているアルゴリズムを説明するための図であり、図4は、ディスクアレイコントローラが転送先を決定するために行なう動作の流れを示した流れ図であり、図5は、その動作の結果、決定される転送先と転送元の関係を示す模式図である。

【0021】まず、図3を用いて、転送先の決定手順の概要を説明する。前述したように、あるディスク装置のバッファエリアに格納されたデータブロックは、そのディスクを除いたディスク装置が転送先が選ばれる。また、格納エリア内の同一アドレスで指定されるデータブロック中の1つのデータブロックは、パリティの記憶のために使用されるので、“ディスク装置の台数-1”個のデータごとに、格納エリア内の同一アドレスで指定されるデータブロックから転送先を選択してやればよいことになる。

【0022】たとえば、図3(a)のように、“0”から“11”と番号付けられた12個のデータが、ディスク装置のバッファエリアに格納されていた場合、番号順に3つつつデータブロックを抽出して、各データブロックが存在する列(すなわち、ディスク装置)を変更することなく、それぞれの3つのデータブロックを同一行に並べると図3(b)のようになる。たとえば、(b)の2行目のデータブロックが描かれていない部分が一番右端に来るように、この行を回転させてやれば、(c)に示したように、各データを、元の列とは違う列に移動することができる。3行目および4行目も同様の処理により、その行に存在する各データを、元の列と違う列に配置できることになる。

【0023】このような変換により得られる配置は、上述の条件を満たすものであり、後ほど説明するディスクアレイコントローラによる転送先の決定は、基本的には、このようなアルゴリズムで行なわれている。ただ

し、図3(a)の1行目だけは、データブロックが描かれていない部分が、既に、右端の列に存在しているので、上述のものとは異なるアルゴリズムにより、その転送先は決定される。

【0024】図4を用いて、ディスクアレイコントローラが実際に行なう処理を説明する。この処理は、バッファエリア内のj番のデータブロックの転送先となる格納エリア内のデータブロックの番号を求めるものであり、得られたブロック番号は、変数A_jに記憶される。なお、ブロック番号jは、U×(L-1)で規定され、Uは、そのデータブロックが格納されたディスク装置の番号(0~3)であり、Lは、それぞれのエリア内における、そのデータブロックが格納順であり、第1の実施例のディスクアレイ装置では、このブロック番号が、バッファエリアおよび格納エリアで独立に規定されている。

【0025】ディスクアレイコントローラは、まず、カウンタN_{BUFF}、N_{DATA}に、それぞれ、“0”をセット(ステップS201)し、次いで、iとjに、それぞれ、“0”と“N_{BUFF}+U_{MAX}”をセット(ステップS202)する。ここで、U_{MAX}は、ディスク装置番号の最大値であり、ディスクアレイ装置が内蔵するディスク装置台数から“1”を減じた値である。また、カウンタN_{BUFF}は、バッファエリア内の処理対象とする一連のブロックの最初のブロック番号を記憶するためのカウンタであり、N_{DATA}は、格納エリア内の転送先となる一連のデータブロックの最初のブロック番号の記憶するためのカウンタである。

【0026】そして、j番のデータブロックが格納されているディスク装置番号であるU_jと、U_{MAX}との比較を行い、j番のデータブロックがU_{MAX}番のディスク装置に格納されたものでない場合(ステップS203; N)には、“N_{BUFF}+i”番のデータブロックの転送先を、格納エリア内の“N_{DATA}+i”番のデータブロックに決定(ステップS204)する。そして、iに“1”を加算(ステップS205)し、iがU_{MAX}と一致していなければ(ステップS206; N)、ステップS204に戻り、iがU_{MAX}と一致しているときには(ステップS206; Y)、ステップS211に進む。

【0027】また、U_{MAX}番のディスク装置に格納されたものである場合(ステップS203; Y)には、“N_{BUFF}+i”番のデータブロックの転送先を、格納エリア内の“N_{DATA}+i+1”番のデータブロックに決定(ステップS207)する。そして、iに“1”を加算(ステップS208)し、iが“U_{MAX}-1”と一致していなければ(ステップS209; N)、ステップS207に戻る。そして、iが“U_{MAX}-1”と一致しているときには(ステップS209; Y)、“j+1”番のデータブロックの転送先を、格納エリア内の“N_{DATA}”番のデータブロックに決定(ステップS210)して、ステップS211に進む。なお、このステップS207ない

しステップS210で行なっている処理が、図3の一行目に対する処理である。

【0028】ステップS211では、 N_{DATA} 、 N_{BUFF} にそれぞれ、“ $U_{MAX} + 1$ ”、 U_{MAX} が加算され、 N_{BUFF} が N_{CHST} と一致していなければ（ステップS212；N）、ステップS202に戻る。そして、 N_{BUFF} と N_{CHST} が一致したときに（ステップS212；Y）、転送先の設定処理は終了する。

【0029】この一連の処理により、図5に示したように、4台のディスク装置のバッファエリアに記憶されているデータブロックのそれぞれに、3台のディスク装置の格納エリア内のデータブロックが格納先として指定される。図5において、それぞれのデータブロック内に表記した数字は、データを識別するために付したデータ番号であり、ディスク装置14₃の格納エリアに表記してある P_{L-M} は、L番からM番のまでのデータを基に算出されるパリティを示す。なお、データ番号は、バッファエリア内では、ブロック番号と一致している。

【0030】第1の実施例のディスクアレイ装置では、このように、データの転送先が他のディスク装置に指定されるので、データを読み出しながら書き込むといった動作が可能となり、短い時間でデータ転送を完了することができる。しかし、データの転送を行なう順番が適当なものでないと、各ブロックのデータ転送の完了後に、ヘッドが長い距離をシークしなければならない。たとえば、データ番号順に、データ転送を行なわせると、書き込みに用いられていたヘッドが次の処理では読み出しに用いられるといったケースが多く現われる。このため、第1の実施例のディスクアレイ装置では、それぞれのディスク装置に対するデータの書き込みを完了させた後に、次のディスク装置に対する書き込みを実行させるといった手順で、データ転送を行なっている。

【0031】すなわち、図5の状態に対して、データ転送を行なう場合には、まず、ディスク装置14₀（ $U=0$ ）内の格納エリアに対するデータ転送を全て完了させてしまう。このとき、最後の“データ9”の転送を開始する前に、ディスク装置14₂のヘッドを格納エリア側に移動させておく。そして、ディスク装置14₀へのデータ転送の完了後、ディスク装置14₂へのデータ転送を開始させる。

【0032】このような順で、データ転送を実行することにより、第1の実施例のディスクアレイ装置では、データ転送に必要とされる時間を、更に短縮化している。

【0033】なお、以上の説明においては、RAIDレベル4の冗長構成を形成させる場合を例としたが、第1の実施例のディスクアレイ装置では、RAIDレベル5の冗長構成でデータを記憶させることもできる。この場合には、図3（b）に示した状態を、右または左に“1”ずつ、ずらしてやれば、それぞれのデータに対して、必要な条件を満たす転送先を設定してやることがで

きる。参考として、図6に、RAIDレベル5の冗長構成を形成する場合の、ディスクアレイコントローラの動作の流れを示し、図7に、その結果として得られるデータブロックの配置を示しておく。

【0034】図6に示した流れは、図4と類似するものである。その概要だけを説明することにする。カウンタ N_{BUFF} 、 N_{DATA} には、処理対象となるバッファエリア、格納エリアの最初のブロック番号が入り、各ループ（ステップS302ないしS309）においては、バッファエリア内に格納されている N_{BUFF} 番から“ $N_{BUFF} + U_{MAX} - 1$ ”番までのデータブロックの転送先が決定されている。各データブロックの転送先は、それぞれのデータブロックが格納されたディスク装置の右隣のディスク装置に設定（ステップS303ないしステップS308）している。当然、右端のディスク装置内のデータブロックの転送先は、左端のディスク装置内に設定（ステップS306）される。

【0035】変形例

【0036】第1の実施例のディスクアレイ装置では、3ブロック分のデータを読み出して、パリティを作成し、作成したパリティを書き込むといった動作が必要であるが、パリティ生成回路と各ディスク装置とを1本のバスで接続していたため、パリティ作成処理時にある程度の時間が必要とされる。このパリティ作成処理時間を短縮するため、変形例のディスクアレイ装置では、3つのデータを同時に処理できるパリティ生成回路と、バス切替回路を用いて装置を構成した。

【0037】図8に、変形例のディスクアレイ装置の構成を示す。変形例のディスクアレイ装置11は、ディスクアレイコントローラ12とパリティ生成回路13と4台のディスク装置14とバス切替回路18で構成される。

【0038】図9を用いて、バス切替回路18およびパリティ生成回路13の動作の説明を行なう。バス切替回路18は、通常時には、各ディスク装置14をディスクアレイコントローラ12と接続する。そして、ディスクアレイコントローラから、パリティの作成を指示されたときには、図9に示したように、内部の接続を切り替え、各ディスク装置と直接にデータのやり取りができる経路を構成する。そして、3台のディスク装置14₀ないし14₂から、パリティを算出するデータを同時に読み出し、それらの排他的論理和演算を行い、その演算結果であるパリティをディスク装置14₃に書き込む。

【0039】このように、変形例のディスクアレイ装置では、第1の実施例のディスクアレイ装置に比べて、パリティ算出時に必要とされる時間が少なくなるため、総合的な動作時間の短縮が図れる。

【0040】第2の実施例

【0041】図10に、第2の実施例のディスクアレイ装置の構成を示す。第2の実施例のディスクアレイ装置

は、回路的には、第1の実施例のディスクアレイ装置に圧縮伸長回路19を付加した構成をとる。圧縮伸長回路19は、ホストコンピュータ21等の外部装置から送られてきたデータの圧縮、および、ディスク装置内に格納された圧縮データの伸長を行なう回路であり、圧縮時には、圧縮率の算出も同時に行なう回路である。このディスクアレイ装置では、それぞれのディスク装置内に、転送レートに応じて分類された記憶領域である高転送レート領域22と中転送レート領域23と低転送レート領域24が設定され、書き込み要求のなされたデータは、圧縮された後に、その圧縮率に応じた転送レート領域に格納される。

【0042】図11に、第2の実施例のディスクアレイ装置における、ディスクアレイコントローラのデータ格納時の動作の流れを示す。データの書き込み要求を受けたディスクアレイコントローラは、そのデータの圧縮率に応じて、データを格納する記憶領域を選択する（ステップS401ないしステップS405）。これらのステップにより、圧縮率（圧縮後のデータサイズ/圧縮前のデータサイズ）が定数C₁より小さいデータに対しては、低転送レート領域に格納することが指定（ステップS402）され、圧縮率が定数C₁以上、C₂未満のデータには、中転送レート領域が指定される（ステップS404）。また、圧縮率が定数C₂以上のデータの格納領域としては、高転送レート領域が指定される（ステップS405）。

【0043】その後、その指定された記憶領域にそのデータの格納に十分な空き領域があるか否かが判断（ステップS406）され、空き領域があった場合（Y）には、そこにデータを格納（ステップS408）する。圧縮率を基に選択された記憶領域に空き領域が無かった場合（ステップS406；N）には、他の記憶領域の空き領域を探してそこに書き込む（ステップS407）。なお、ディスクアレイ装置内にそのデータを記憶する空き容量がない場合には、ここに示した処理に入る前に、エラーメッセージが出力される。

【0044】このように、このディスクアレイ装置では、圧縮率が悪いデータは、転送レートが高い記憶領域に格納され、逆に、圧縮率が良いデータは、転送レートが低い記憶領域に格納される。このため、圧縮率が悪いデータが転送レートの低い記憶領域に格納されるようなことがなくなり、ディスク装置としての総合的な応答速度を向上させることができる。

【0045】第2の実施例のディスクアレイ装置では、外部装置から、圧縮されていないデータを受けて、装置内で圧縮を行なうように構成してあるが、外部装置から圧縮されたデータとその圧縮率を受けるように構成してもよい。また、ステップS406およびS407の書き込みにおいては、RAIDレベル4のようなパリティを用いた構成をとるようにしてもよく、RAIDレベル0

のようなパリティを用いない構成をとるようにしてもよい。また、パリティを使用する際には、変形例のディスクアレイ装置のように、バス切替回路を組み合わせ、各ディスク装置内のデータをパリティ生成回路が同時にアクセスできるように構成してもよい。

【0046】

【発明の効果】以上説明したように、請求項1記載の発明では、複数のディスク装置内の高速にデータ転送が行なえる記憶領域をバッファとして用いるため、大量のデータを一時に受信した場合の動作速度は、従来のディスクアレイ装置よりも早くなる。また、異なるディスク装置間のデータ転送だけで冗長構成を形成するように構成されているので、高速に冗長構成を形成することができる。

【0047】請求項2記載の発明では、圧縮率が悪いデータは、転送レートが高い記憶領域に格納し、逆に、圧縮率が良いデータは、転送レートが低い記憶領域に格納する。このため、圧縮率が悪いデータが転送レートの低い記憶領域に格納されるようなことがなくなり、ディスク装置としての総合的な応答速度を向上させることができる。

【図面の簡単な説明】

【図1】 本発明の第1の実施例によるアレイ型ディスク装置の概要を示す構成図である。

【図2】 第1の実施例のアレイ型ディスク装置における、バッファエリアから格納エリアへのデータ転送の流れを示す流れ図である。

【図3】 第1の実施例のアレイ型ディスク装置における、データ転送先の決定手順を説明するための説明図である。

【図4】 第1の実施例のアレイ型ディスク装置における、転送先の決定動作の流れを示す流れ図である。

【図5】 第1の実施例のアレイ型ディスク装置における、転送先の決定結果の一例を示す説明図である。

【図6】 第1の実施例のアレイ型ディスク装置においてRAIDレベル5の冗長構成を採用する際の、転送先の決定動作の流れを示す流れ図である。

【図7】 第1の実施例のアレイ型ディスク装置においてRAIDレベル5の冗長構成を採用した際の、転送先の決定結果の一例を示す説明図である。

【図8】 変形例のアレイ型ディスク装置の概要を示す構成図である。

【図9】 変形例のアレイ型ディスク装置の動作内容を示すための説明図である。

【図10】 本発明の第2の実施例によるアレイ型ディスク装置の概要を示す構成図である。

【図11】 第2の実施例のアレイ型ディスク装置における、データ格納動作の流れを示す流れ図である。

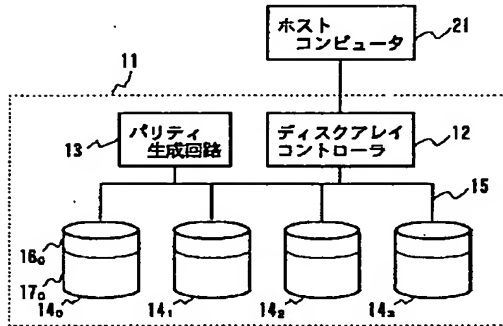
【符号の説明】

11…アレイ型ディスク装置、12…ディスクアレイコ

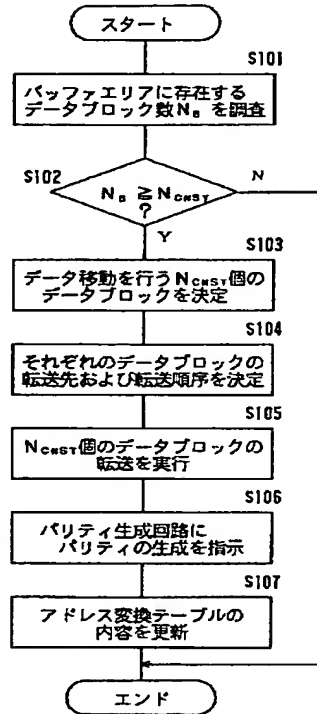
ントローラ、13…パリティ生成回路、14…ディスク装置、15…バス、16…バッファ領域、17…格納領域、18…バス切替回路、19…圧縮伸長回路、21…

ホストコンピュータ、22…高転送レート領域、23…中転送レート領域、24…低転送レート領域

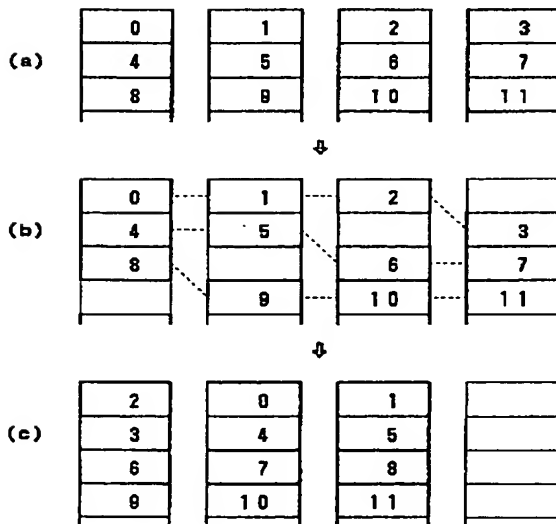
【図1】



【図2】



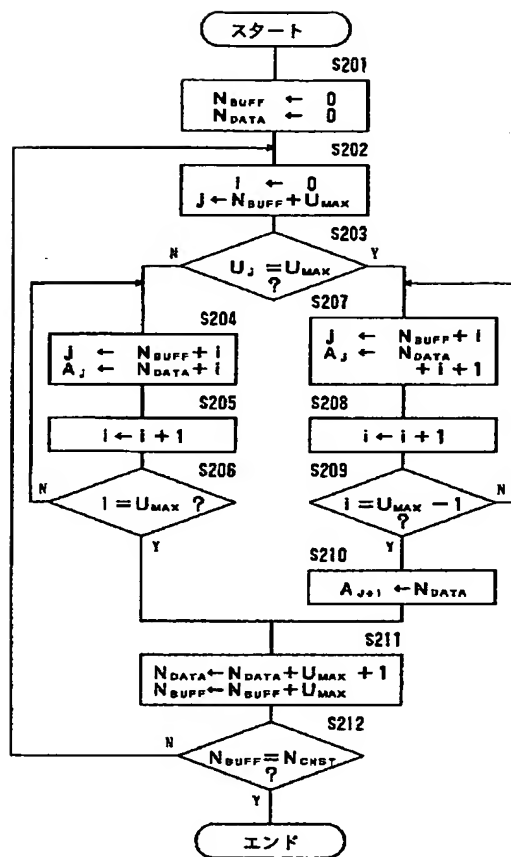
【図3】



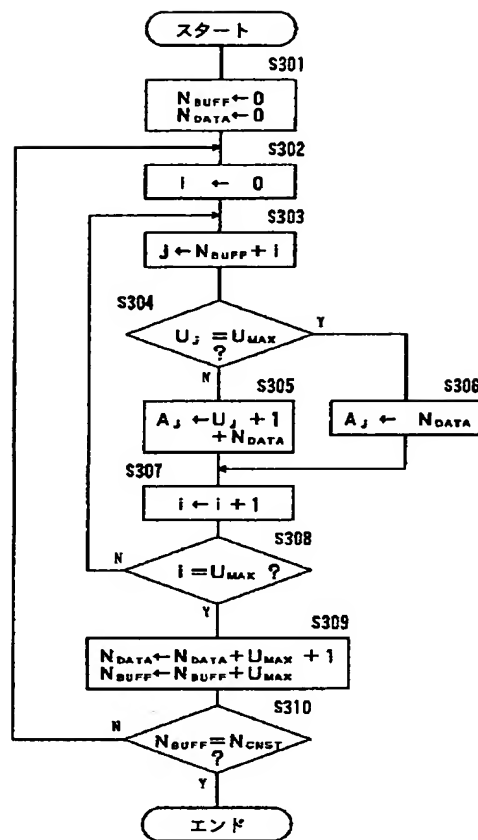
【図5】

		U			
L		0	1	2	3
		0	1	2	3
バッファエリア	0	0	1	2	3
	1	4	5	6	7
	2	8	9	10	11
	⋮				
格納エリア	0	2	0	1	P ₀₋₂
	1	3	4	5	P ₀₋₅
	2	6	7	8	P ₀₋₈
	3	9	10	11	P ₀₋₁₁
	⋮				

【図4】



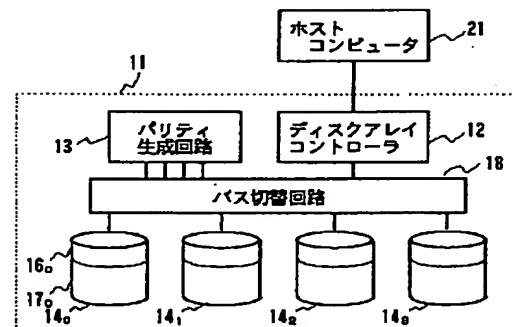
【図6】



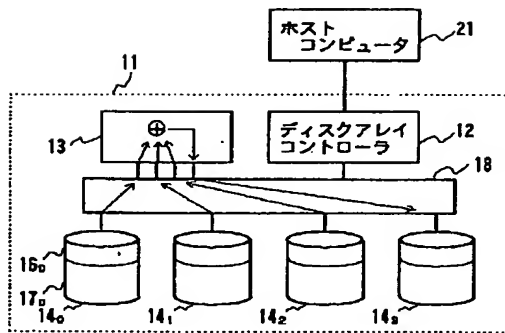
【図7】

L \ U		0	1	2	3
パナフレア	0	0	1	2	3
	1	4	5	8	7
	2	8	9	10	11
	...				
格納エリア	0	P ₀₋₂	0	1	2
	1	3	4	5	P ₃₋₅
	2	7	8	P ₈₋₉	8
	3	11	P ₉₋₁₁	8	10
	...				

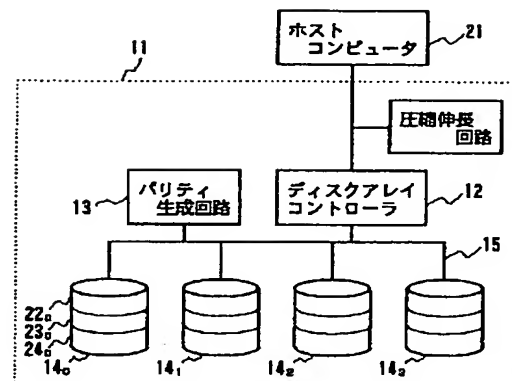
【図8】



【図 9】



【図 10】



【図 11】

